

Dear Editors and Authors,

Sorry, as I don't have enough time to carefully go through the work, specific comments are only focused on the Introduction part of the paper.

General Comments

This work encloses a lot of contents and seems interesting to assess the ET model having the best performances in terms of the criteria the authors select. Language issues should be fully checked throughout the entire text before publication in HESS. Novelty of the paper should be better emphasized rather than "BME has not been used for evaluating the ET models". Model complexity for each model should be better described. For example, authors can directly introduce number of parameters with uncertainties in their experiment? According to my comments, a minor revision is suggested.

Specific comments

1. The Abstract is out of organization. It seems to me that you never mention the model complexity but always write "underestimation" or "overestimate" to explain why the SW is the best one.
2. Lines 25-27: unclear, please rephrase this sentence
3. It is unclear for me why 'SW' is best one from the abstract.
4. Line 37: please check the symbols
5. Simulate ET or estimate ET? Please be very sure of this word.
6. Line 41: add a reference
7. Lines 55-56: unclear, please rephrase this sentence.
8. Lines 62-63: 'These quantitative criteria' refer to what?
9. Line 70: performances
10. Line 71: remove 'the' from 'the SW model'
11. Lines 71-72: please rephrase this sentence
12. Line 73: should be model ranking? Please check the terminology.
13. Lines 75-76: unclear, significant variability of model performances?
14. Lines 92-93: been?
15. Lines 102-103: add a reference